
Die UWE-Befragung: Konzept und Perspektiven der Berichtslegung.

Katharina Knüttel

1 Einleitung

Mit UWE (kurz für Umwelt, Wohlbefinden und Entwicklung von Kindern und Jugendlichen) wird das Wohlbefinden von Kindern und Jugendlichen und ihre Sicht auf ihre Umwelt mithilfe eines standardisierten Fragebogens erhoben. Das bisherige Transferkonzept sieht neben den Stadtberichten für die kommunalen Akteur*innen auch vor, dass jede teilnehmende Schule einen eigenen Schulbericht erhält (Petermann et al. 2019: 37f.). Die dort enthaltenen schulspezifischen Angaben können zur internen Schulentwicklung und zur Förderung einer schulbezogenen Gesprächskultur zwischen Schulleitung, Lehrer*innen, Schüler*innen und ihren Familien sowie weiteren (z.B. sozialpädagogischen) Akteur*innen genutzt werden.

Bislang wurden die Grafiken und Tabellen der Schulberichte zwar syntaxbasiert, reproduzierbar und über alle Schulen hinweg über die Programmierung entsprechender Auswertungsschleifen in der Statistiksoftware Stata produziert, die anschließende Berichterstattung erfolgte dann aber über die manuelle Integration der Grafiken und Tabellen in ein Textverarbeitungsprogramm. Im Hinblick auf die Skalierung der Befragung ist es sinnvoll, die Berichtslegung weiter zu standardisieren und zu automatisieren. In den Pilotphasen war die manuelle Integration der Ergebnisse in Textdokumente noch gut leistbar und der Aufwand der Erstellung einer automatisierten Berichtslegung höher als der des manuellen Vorgehens. Bei einer Ausweitung des Konzepts auf weitere Kommunen oder der weiteren Wiederholung der Befragung dreht sich dieses Verhältnis allerdings um.

Im Folgenden wird daher eine Möglichkeit der automatisierten Berichtslegung mit R, RStudio und RMarkdown vorgestellt. Im Rahmen der Fortführung des Projekts für die Ruhr-Konferenz ([Link zur Projektbeschreibung](#)) befindet sich der erste Teil des Konzepts bereits in der Erprobung und vorläufige Berichtsentwürfe liegen vor. Als weitere Perspektive wird die Entwicklung interaktiver Onlineformate („Shiny Apps“/„Dashboards“) vorgeschlagen. Hier können die Schulen sich mit ihren Zugangsdaten einloggen und sich ihre Ergebnisse, auch für den Zeitvergleich, selbst zusammenstellen und herunterladen.

2 Ein Konzept für automatisierte Berichtslegung

In der vorliegenden Darstellung werden zunächst die Schulberichte skizziert (2.1), bevor knapp erläutert wird, was sich hinter R, RStudio und RMarkdown eigentlich verbirgt (2.2). Anschließend wird das Verfahren erläutert und es werden Beispiele aus der aktuellen Berichtslegung gezeigt (2.3), bevor mögliche Anschlussperspektiven interaktiver Berichterstattungsformate für die Schulen skizziert werden (2.4).

2.1 Schulberichte

Die Schulberichte sind ein zentrales Element in der UWE-Befragung: Nur, wenn in den Schulen bekannt ist, wie es den Kindern und Jugendlichen dort geht und über welche Ressourcen sie in welchem Ausmaß verfügen, kann auch ein Transfer dieser Erkenntnisse gelingen und ein Prozess gestartet werden, in dem gezielte Maßnahmen zur Förderung von Wohlbefinden und Ressourcen der Kinder und Jugendlichen initiiert werden. Jede Schule erhält daher einen eigenen, nicht öffentlichen Bericht, in dem die schulspezifischen Befragungsergebnisse dargestellt sind. Die Schulberichte sind alle gleich aufgebaut und bestehen aus

- einer allgemeinen Einführung in die Befragung und die zentralen Themenbereiche
- allgemeinen Hinweisen zur Interpretation der Tabellen und Grafiken
- schulspezifischen Auswertungen in tabellarischer und grafischer Form, zum Beispiel Anteilswerte, Mittelwerte und Balkendiagramme.

Die Auswertung erfolgt immer für die jeweilige Schule im Vergleich zum gesamtstädtischen Wert. Für Schule A dann beispielsweise: An Ihrer Schule A haben 80% der befragten Kinder und Jugendlichen einen Migrationshintergrund, insgesamt haben in Kommune C 75% der befragten Kinder und Jugendlichen einen Migrationshintergrund. Die Schulen bekommen damit Ergebnisse für ihre Kinder und Jugendlichen und zusätzlich zu der reinen Beschreibung dieser Ergebnisse eine Einschätzung dazu, wie die Schüler*innen an ihrer Schule im Vergleich zu den anderen Kindern und Jugendlichen hinsichtlich des Wohlbefindens und ihrer Ressourcen aufgestellt sind.

2.2 Bestandteile der automatisierten Berichtslegung (Software/Programmiersprachen)

Eine Möglichkeit, die Automatisierung durchzuführen, besteht in der Verwendung von R, RStudio und der RMarkdown. Im Folgenden werden diese drei Bausteine kurz erläutert, bevor das Vorgehen für die automatisierte Berichtslegung und sinnvolle Weiterentwicklungen skizziert werden.

R

R ist eine statistik-orientierte Programmiersprache. R wird häufig als Statistik-Software beschrieben, allerdings ist es im Verständnis der verantwortlichen Entwickler*innen eher eine „Umgebung“. Dieser Begriff soll betonen, dass es sich um ein umfassend geplantes kohärentes System handelt – statt einer Ansammlung spezifischer und unflexibler Tools, wie sie es für andere Datenanalysesoftware diagnostizieren (Venables et al. 2020: 2). Unabhängig von der Bezeichnung gilt: R ist ein weltweit genutztes Instrument, um statistische Auswertungen vornehmen zu können. Es ist kostenlos und „open source“ und es gibt eine international tätige Gemeinschaft, die ständig daran arbeitet, Zusatzfunktionen in Form von „packages“ zu erarbeiten und zur Verfügung zu stellen. Die mittlerweile verfügbaren Pakete umfassen ein breites Spektrum fortgeschrittener Verfahren, von Geodatenanalyse bis hin zu Strukturgleichungsmodellen. Es gibt wenige Verfahren, die sich mit R nicht umsetzen ließen, und das wachsende Angebot an „packages“ verringert stetig den Anteil an nicht umsetzbaren Methoden.

RStudio

R selbst ist – wie beschrieben – eine Programmiersprache. Für mehr Komfort und Funktionen bei der Programmierung nutzen Anwender*innen in der Regel eine zusätzliche Entwicklungsumgebung/grafische Oberfläche. Eine davon, die auch in einschlägigen aktuellen Lehrbüchern genannt wird (z.B. Luhmann 2020, Wollschläger 2020) und viele zusätzliche Möglichkeiten bietet, ist RStudio. Einige der Vorzüge dieser Software sind die Möglichkeiten, Text und statistische Auswertungen zu kombinieren und so syntaxbasiert z.B. PDF-, HTML- oder MS-Word-Outputs zu generieren. Weitere Vorteile liegen in der großen Entwicklungsdynamik der letzten Jahre, zum Beispiel hinsichtlich dynamischer Darstellungsmöglichkeiten (Anwendungsbereiche sind hier z.B. Dashboards oder Online-Tutorials).

Markdown/RMarkdown

Markdown ist eine sogenannte Auszeichnungssprache (und der Name dementsprechend ein Wortspiel – Auszeichnungssprache heißt auf Englisch „Markup Language“). Das bedeutet, dass die Textbestandteile entsprechend ihrer Eigenschaften gekennzeichnet werden: Handelt es sich um eine Überschrift, einen Absatz, eine nummerierte Liste etc.? Davon zu unterscheiden ist das vom Team von RStudio entwickelte und dort bereits integrierte „RMarkdown“. Mit RMarkdown wird die Erstellung von Textdokumenten mit integrierten statistischen Analysen mit Hilfe von Auszeichnungssprachen wie Markdown ermöglicht. Ein solches Vorgehen ist vorteilhaft, da die „manuelle“ Übertragung von Ergebnissen aus der Statistiksoftware in Textdokumente grundsätzlich fehleranfälliger ist. RMarkdown verbreitet sich seit seiner Entwicklung rasant und mittlerweile werden mit dem zusätzlichen Paket „bookdown“ ganze Bücher damit publiziert (<https://bookdown.org/>).

2.3 Berichtslegung

Die eigentliche Berichtslegung erfolgt dann als „parameterized report“ (Xie et al. 2020: Kapitel 15): In einem RMarkdown-Dokument werden sowohl die über Markdown oder z.B. LaTeX (LaTeX ist ebenfalls eine Auszeichnungssprache) ausgezeichneten Textbausteine als auch die mit R programmierten Auswertungen angelegt. Die Analysen beziehen dabei einen schulspezifischen „Parameter“ wie die Schulnummer oder den Schulnamen ein. Die Programmierlogik lautet dann z.B.: „Erstelle für alle Schulnummern einen eigenen Bericht mit den vorab definierten Analysen. Wähle für den Vergleich jeweils die Kinder der ausgewählten Schule im Vergleich zu allen Kindern“. Die erste Programmierung der Berichte ist dabei relativ aufwändig, da für jede Grafik und Bedingung einige Zeilen Programmiercode zu entwickeln sind und auch die Textbausteine mit LaTeX oder Markdown ausgezeichnet sein müssen. Wenn dieser Code jedoch einmal entwickelt ist, können damit innerhalb eines Tages mehrere Hundert Schulberichte generiert werden. Damit ist die automatisierte Berichtslegung im Wesentlichen abgeschlossen und die Schulberichte können den Schulen zugestellt werden.

Dieses Vorgehen wird derzeit in der Berichtslegung der UWE-Weiterentwicklung für die Ruhr-Konferenz erprobt. Die folgenden Abbildungen sollen die Möglichkeiten des Vorgehens mit Beispielen aus der aktuellen Berichtslegung der UWE-Grundschulbefragung in Herne veranschaulichen. Abbildung 1 zeigt einen Kartentyp, der in den Schulberichten vorzufinden ist. Um das räumliche Umfeld – und damit auch größtenteils das Einzugsgebiet – der jeweiligen Grundschule einordnen zu können, wird in jedem Schulbericht zunächst die gleiche Kartengrundlage verwendet. Es handelt sich um die statistischen Bezirke in Herne, die entsprechend der Höhe der Kinderarmutsquote eingefärbt sind.¹ In jedem Schulbericht wird nun automatisch der Schulstandort der jeweiligen Schule als Punkt („Ihre Schule“) erzeugt.²

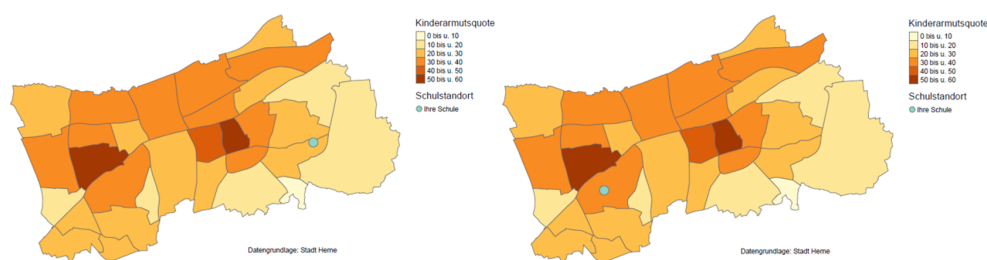


Abbildung 1: Beispiele für die Kartierung von Schulstandorten aus zwei Schulberichten

¹Kinderarmutsquote: Anteil Kinder und Jugendliche unter 15 Jahren, die Leistungen nach dem SGB II beziehen, an allen Kindern und Jugendlichen unter 15 Jahren im Dezember 2019. Datengrundlage: Stadt Herne

²Da die beiden abgebildeten Schulen anhand ihres Standorts zu identifizieren sind, beziehen sich die anderen Auszüge, die auch Befragungsergebnisse beinhalten, auf andere als die hier kartierten Schulen.

Ebenfalls automatisiert können Ergebnisse in Fließtext eingebaut werden (Abbildung 2 oder z.B. Balkendiagramme erzeugt werden (Abbildung 3). Der gleiche Code erzeugt dann so viele Schulberichte mit den exemplarisch in den Abbildungen dargestellten Auswertungsarten, wie es erforderlich ist. In jedem Bericht werden automatisch die richtigen Werte der jeweiligen Schule eingefügt. Übertragungsfehler durch Copy & Paste in Textdokumente sind dabei ausgeschlossen.

Außerdem wurden verschiedene Aussagen zur Beziehungsqualität zu Erwachsenen gestellt, von denen wir für die Schulberichte eine als besonders relevant hervorheben möchten: Die Frage danach, ob es für die Kinder an ihrer Schule mindestens eine erwachsene Person gibt, der sie „wirklich wichtig“ sind. Die Schulen in Herne unterscheiden sich erheblich in den Anteilen der Kinder, die hier zustimmen.

An Ihrer Schule sind das **76%**. An allen Grundschulen in Herne liegt der Anteil bei **63%**.

Abbildung 2: Auszug aus einem Schulbericht: Automatisierte Berechnung im Fließtext

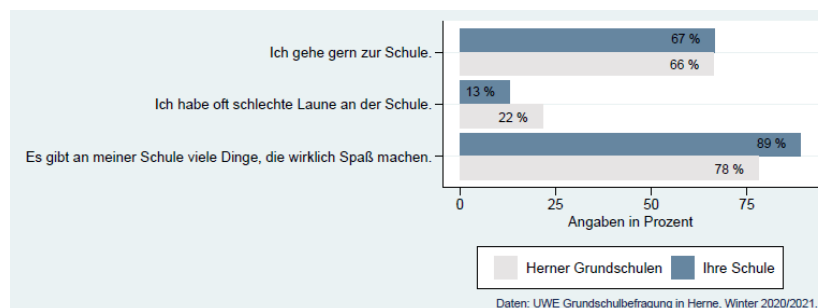


Abbildung 3: Beispiel für ein Balkendiagramm aus einem Schulbericht

Zuletzt sei noch ein weiterer Vorteil dieser Form der Berichtslegung genannt: In RMarkdown-Dokumenten ist es ebenfalls möglich, die Darstellung bestimmter Bausteine an Bedingungen zu knüpfen. Das ist vor allem deswegen interessant, weil es je nach Auswertungsanliegen zu Fallzahlproblemen kommen kann – es ist dann unter Umständen sinnvoll, auf bestimmte Darstellungen zu verzichten oder Warnhinweise einzublenden. Abbildung 4 zeigt eine Seite aus einem Schulbericht. Nur die Schulen, an denen weniger als 30 Schülerinnen und Schüler an der Befragung teilgenommen haben, erhalten den vorab sichtbaren Warnhinweis zur Belastbarkeit der Aussagen. Außerdem wird die Darstellung der Ergebnisse im kompletten Schulbericht angepasst: Statt der exakten Anteilswerte werden den Schulen mit geringen Fallzahlen Informationen darüber gegeben, ob sie vom städtischen Durchschnittswert stark abweichen – und wenn ja, in welche Richtung.

Rücklauf und soziodemografische Angaben

Vorbemerkung

Achtung: Für Ihre Schule liegen weniger als 30 ausgefüllte Fragebögen vor. Statistiken mit so wenig Fällen sind nur bedingt aussagekräftig. Wenn von z.B. zwanzig Kindern nur eines eine unerwartete Antwort gibt, entspricht dies bereits einem Anteil von 5%. Wir haben uns deswegen dazu entschieden, in den Diagrammen und Tabellen die *gesamstädtischen* Ergebnisse darzustellen.

- Wenn keine weitere Bemerkung folgt, bedeutet das, dass die Anteilswerte an Ihrer Schule sich nahe dem gesamstädtischen Werte ($\pm 10\%$) bewegen.
- Wenn an Ihrer Schule der Anteilswert um mehr als 10 Prozentpunkte abweicht, wird dies unter der jeweiligen Auswertung ausgewiesen und *kann* auf eine besondere Lage an der Schule hindeuten.

Abbildung 4: Auszug aus einem Schulbericht mit Warnhinweis bei geringen Fallzahlen

2.4 Perspektive: Interaktive Darstellung über Shiny Apps

Die standardisierten Schulberichte liefern einen Überblick über die Befragungsergebnisse der jeweiligen Schule im Vergleich mit der Gesamtheit der Schulen ihrer Stadt. Allerdings muss dabei aus allen möglichen Darstellungsformen eine gewählt werden, und schon die reine Beschreibung der Ergebnisse führt zu Schulberichten im Umfang von knapp 50 Seiten. Wenn an einer Schule gezielt Kinder gefördert werden sollen, ist es aber unter Umständen sinnvoll, genauer in die Daten zu schauen: Sind es z.B. vielleicht vor allem Mädchen, die ein schlechtes Körperbild haben? Welche Kinder erfahren häufiger Mobbing? Sämtliche möglichen interessanten Zusammenhänge in die Schulberichte zu integrieren, würde diese enorm überladen. Es gibt aber die Möglichkeit, einem Zielpublikum flexible Darstellungen zu erlauben. Dazu können sogenannte „Dashboards“ entwickelt werden: Interaktive digitale Berichterstattungsformate.

Ein Dashboard ist eine grafische Oberfläche, die zur Visualisierung von Daten dient. Die Anwender* können per Mausclick bestimmte Themenfelder, Indikatoren und Darstellungsformen anwählen. Es besteht die Möglichkeit, einen Grafik- und Tabellenexport zu integrieren.

Das Grundgerüst eines solchen Dashboards kann mit Hilfe von R, RStudio und RMarkdown angelegt werden, allerdings werden zwei zusätzliche Elemente benötigt:

1. Shiny Apps: Die interaktive Darstellung erfolgt dann über sogenannte „Shiny Apps“, in denen die Auswahl von Indikatoren über grafische Elemente wie Dropdownmenüs und die Struktur der Darstellung über ein Menü und/oder verschiedene „Reiter“ angelegt werden muss. Die Möglichkeiten zur Programmierung solcher „Shiny Apps“ sind über die weiter oben beschriebenen „packages“ in R und RStudio kostenlos möglich.
2. RStudio Server: Der digitale Zugang muss über einen RStudio-Server erfolgen. Die Sensibilität schulbezogener Angaben setzt die Installation eines eigenen RStudio-Servers voraus. Jede Schule erhält dann eigene Zugangsdaten (Benutzernamen und Passwort) und hat nur Zugriff auf ihre eigene Datendarstellung. Der Zugang erfolgt über den Browser und setzt keine speziellen Kenntnisse voraus.

Der sinnvolle Aufbau eines solchen Dashboards benötigt allerdings Zeit und muss sorgfältig programmiert werden. Es muss im Vorfeld festgelegt werden, welche Darstellungsformen ermöglicht werden und wie das Dashboard strukturiert wird. Zur sinnvollen Interpretation müssen zusätzliche Informationen zur Darstellung formuliert und angezeigt werden. Gerade wenn es um die Darstellung der Ergebnisse einzelner Fragen an einzelnen Schulen geht, muss man hinsichtlich der Belastbarkeit auch an einigen Stellen überlegen, ob bei geringen Fallzahlen entsprechende Warnhinweise eingeblendet werden und ob Ergebnisse unterhalb einer bestimmten Fallzahl überhaupt dargestellt werden sollen.

Es wäre zudem möglich, Zeitvergleiche zu integrieren. Die UWE-Befragung ist als Bestandteil eines Monitoringprozesses angelegt, das heißt, dass die Daten fortlaufend über mehrere Zeitpunkte hinweg beobachtet werden sollten. So haben die weiterführenden Schulen der Pilotkommune Herne mittlerweile an zwei Erhebungswellen teilgenommen. Um Entwicklungen an den Schulen zu betrachten, können natürlich die Schulberichte jeweils nebeneinandergelegt und abgeglichen werden, was jedoch ein eher mühseliger Prozess ist. Entweder könnten für alle Schulen, die mehrmals an der Befragung teilnehmen, auch standardisierte Trendberichte erstellt werden – oder aber sie erhalten die Möglichkeit, die Entwicklung innerhalb eines solchen Dashboards zu betrachten.

Ganz gleich, ob es um Zeitvergleiche, vertiefende Betrachtung von Zusammenhängen oder alternative grafische Darstellungsformen geht: Ein gut angelegtes Dashboard ermöglicht eine flexibel an die jeweiligen Interessen und Bedarfe angepasste Darstellung.

3 Abschluss

Im vorliegenden Text wurde ein Konzept für die standardisierte Berichtslegung im Rahmen der UWE-Befragung vorgelegt, welches einerseits die automatisierte Erstellung von Schulberichten in Form von Textdokumenten mit integrierten Karten, Tabellen und Grafiken als auch die interaktive Darstellung über „Shiny Apps“ und Dashboards beinhaltet.

Die Automatisierung der Schulberichte mit R, RStudio und RMarkdown ist derzeit bereits im Rahmen der Grundschulbefragung in der Erprobungsphase. Darauf aufbauend können mit den gleichen Instrumenten und der zusätzlichen Einbindung von „Shiny Apps“ und der Installation eines RStudio Servers auch interaktive, flexible Berichterstattungsformate entwickelt werden. Auch zur Gestaltung von „Shiny Apps“ und Dashboards liegen bereits erste Erfahrungen vor, allerdings erfordert deren Gestaltung einen weitaus größeren Zeitaufwand und damit eine längerfristige Perspektive. Insbesondere aufgrund der Möglichkeit für die Schulen, individualisierte Auswertungen und komfortable und flexible Zeitvergleiche vornehmen zu können, kann sich dieser Aufwand lohnen.

Sämtliche verwendete Software ist kostenlos und Open Source, jedoch wird für die Installation eines RStudio-Servers die kostenpflichtige „Pro-Variante“ empfohlen. Für das Hosting des RStudio-Servers und der Zugangsdaten gibt es bereits erste Erfahrungen in Lehrprojekten der Ruhr-Universität Bochum, ggf. kann es sinnvoll sein, mit Partner*innen aus dem IT-Bereich zu kooperieren.

Die vorliegenden Ausführungen bezogen sich auf die Schulberichte, da hier eine Vielzahl von Berichten nach dem gleichen Schema und mit den gleichen Bausteinen erstellt wird. Im Fall einer Ausweitung des Instruments auf weitere Kommunen kann dieses Vorgehen allerdings selbstverständlich auch für die UWE-Stadtberichte durchgeführt werden. Voraussetzung für die Umsetzung ist allerdings eine zuverlässige Perspektive, denn die Entwicklung eines solchen Dashboards ist nur für die standardisierte Dauerbeobachtung sinnvoll.

Literaturverzeichnis

- Luhmann, Maike (2020): R für Einsteiger. Einführung in die Statistik-Software für die Sozialwissenschaften. 5., überarbeitete Auflage. Weinheim/Basel: Beltz.
- Petermann, Sören; Strohmeier, Klaus Peter; Yaltzis, Katharina; Albrecht, Michaela (2019): Wie geht's dir, UWE? Monitoring von Umwelt, Wohlbefinden und Entwicklung von Kindern und Jugendlichen im Sekundarschulalter als Grundlage einer integrierenden Stadtentwicklung. Ergebnisse der UWE-Pilotstudie 2017 in Herne. FGW-Studie.
- Venables, William N.; Schmith, David M. and the R Core Team (2021): An Introduction to R. Notes on R: A Programming Environment for Data Analysis and Graphics. Version 4.0.5 (2021-03-31). Online hier verfügbar.
- Wollschläger, Daniel (2020): Grundlagen der Datenanalyse mit R. Eine anwendungsorientierte Einführung. 5. Auflage. Berlin: Springer Spektrum.
- Xie, Yihui; Allaire, J.J.; Golemund, Garrett (2020): R Markdown: The Definitive Guide. Boca Raton/London/New York: Chapman and Hall/CRC. Online hier verfügbar.